

Math 401: Lab # 1
Gender Discrimination ¹
Due: January 25, 2016

Adelphi University is a mid-sized university on Long Island, NY. In 2009, they paid \$305,889 to 37 faculty members to settle a pay discrimination lawsuit ². The claimants argued that female faculty of the same or lesser rank were being paid less than their male counterparts. In this lab we will determine, based on the data, whether there was a statistically significant difference.

Question 1. *Given that we are studying salaries for men vs. salaries for women, which statistical inference test should we use?*

The faculty salary data from the Statistics and English departments in 2009 are available as Faculty.txt, in a folder called Lab1, in the Resources part of our Shared Drive.

Question 2. *Without even looking at the data, which gender do you expect to make more? ³ Which department do you expect professors to be paid more?*

Wrangle the data so that you have it in either R or StatKey. Record the steps you took in your Lab Report, following the Lab Report Guidelines. If you discover missing data or outliers, decide what to do about it, and be sure to include a justification in your Lab Report. Discuss as well how your decision might or might not affect your analysis.

Produce the most useful two graphs or charts of the Faculty data that you can to explain the story within in the data. For example: what are the key patterns or trends in the data? You should consider gender differences at every level (assistant, associate, full) and between the two different departments. Include the image you produce and text explaining how it should be interpreted.

Question 3. *What explanatory variables seem to influence salary? What biases or lurking variables could be influencing our question of interest?*

Question 4. *Are there any unusual observations, or outliers, that seem to not fit the rest of the data? Do these observations have the ability to sway our conclusions?*

Carry out the test you decided on in Question 1. Explain why you chose that test, why randomization methods are appropriate (hint: think about sample sizes in the various groups you are considering). Compute and interpret p-values to go along

¹Adapted from *Practicing Statistics*, by Kuiper and Sklar, Pearson

²Article in Garden City News

³For more info, see Perna “Sex and Race Differences in Faculty Rank and Tenure”, *Research in Higher Education* 42 (2001)

with your analysis. You might consider running multiple tests, corresponding to the different subpopulations. You can also use more advanced tools that we haven't learned in this class yet (e.g. ANOVA), as long as you explain the tool you're using in your Methods section.

Question 5. *Discuss the scope of your analysis. Can it be applied to all mid-sized universities? To all departments at Adelphi?*

Discuss any biases that might be affecting your analysis, but that are beyond your control. In other words, state the scope of your analysis as precisely and unambiguously as you can.

Question 6. *Which department, Statistics or English, has more male professors? How does this affect the salary distribution?*

Submission Notes

In this lab, you will work with a randomly assigned partner. You should work with that partner to wrangle the data into a form where you can test it, either using StatKey, Excel, or R. You may conduct your tests and analyses together, but you should write-up separate reports. Please follow the Lab Report Guidelines at all stages. **Do not share the results of your analysis with anyone else**, i.e. don't spoil the lab for them by telling them what to expect their final conclusion(s) to be.

Write your 2-4 page lab report to the Dean, explaining your findings. You may assume the Dean has only an intro stats background, so (s)he is probably not familiar with randomization tests or the importance of model conditions. Clearly state your conclusions. What do your p-values really mean? What do your graphs and outputs mean?